HIGH RESOLUTION SURFACE GEOMETRY AND ALBEDO BY COMBINING LASER ALTIMETRY AND VISIBLE IMAGES

Robin D. Morris, Udo von Toussaint and Peter C. Cheeseman,
NASA Ames Research Center, MS 269-2, Moffett Field, CA 94035
[rdm,udt,cheesem]@email.arc.nasa.gov

KEY WORDS: Bayesian inference; surface geometry; albedo; computer vision;

ABSTRACT

The need for accurate geometric and radiometric information over large areas has become increasingly important. Laser altimetry is one of the key technologies for obtaining this geometric information. However, there are important application areas where the observing platform has its orbit constrained by the other instruments it is carrying, and so the spatial resolution that can be recorded by the laser altimeter is limited. In this paper we show how information recorded by one of the other instruments commonly carried, a high-resolution imaging camera, can be combined with the laser altimeter measurements to give a high resolution estimate both of the surface geometry and its reflectance properties. This estimate has an accuracy unavailable from other interpolation methods. We present the results from combining synthetic laser altimeter measurements on a coarse grid with images generated from a surface model to re-create the surface model.

RÉSUMÉ

Le besoin d'informations géométriques et radiométriques précises couvrant de grandes étendues devient de plus en plus important. L'altimétrie laser est une des technologies principales pour obtenir ces informations géométriques. Cependant, il est des domaines d'application importants où la plateforme d'observation a son orbite contrainte par les autres instruments qu'elle porte, ce qui limite la résolution spatiale qui peut être enregistrée par l'altimètre. Dans cet article nous montrons comment l'information enregistrée par un des autres instruments communément embarqués, une caméra photographique à haute résolution, peut être combinée avec les mesures de l'altimètre laser pour donner une estimation haute résolution à la fois de la géométrie de surface et de ses propriétés de réflectivité. Cette évaluation offre une exactitude inégalée par d'autres méthodes d'interpolation. Nous présentons les résultats obtenus en combinant des mesures synthétiques d'altimètre laser sur une grille grossière avec des images produites à partir d'un modèle de surface pour recréer le modèle de surface.

KURZFASSUNG

Präzise geometrische und radiometrische Informationen über grosse Areale ist zunehmend von Bedeutung. Die Laser Altimetrie ist eine der Schlüsseltechnologien zur Gewinnung dieser Daten. Allerdings ist in wichtigen Anwendungsfällen die Laser Altimetrie Messung durch weitere Instrumente behindert und daher die räumliche Auflösung eingeschränkt. In dieser Veröffentlichung zeigen wir auf wie die von einer hochauflösenden Kamera (einer fast immer installierten Diagnostik) gewonnenen Bilder mit den Daten der Laser Altimetrie kombiniert werden können um eine präzise Bestimmung der Oberflächenform und ihrer Reflektivitätseigenschaften zu ermöglichen. Diese Art der Oberflächenbestimmung erweist sich einer Splineinterpolationen der Laser Altimetriedaten überlegen. Wir zeigen die Ergebnisse der Oberflächenrekonstruktion aus der Kombination von synthetischen, niedrig aufgelösten Laser Altimetriedaten und Bildern.

1 INTRODUCTION

The need for accurate geometric information for a variety of problems has grown rapidly in the last decades. These needs cover a broad field, from monitoring of environmental changes such as the deformation rates of glaciers, to the creation of 3-dimensional digital city models, and the determination of the shapes of asteroids and features on planets. The demands with respect to the required accuracy are steadily increasing (Rees 1990).

Laser altimetry systems have been able to respond to these demands. However, for many applications, only coarse resolution sampling is available. This is especially true for planetary and small body observations, where the sampling of the surface is constrained by the orbit of the sensor, and this orbit is often determined by the other instruments carried by the spacecraft.

These other instruments usually include a high-resolution optical imager. These images have been previously used to infer a surface reconstruction, solving this inverse problem using Bayesian probability theory (Smelyanskiy 2000, Morris 2001). The accuracy of the reconstruction of the 3-dimensional surface depends on the geometric information content of the images and on additional prior knowledge. Often images from mapping orbits do not contain much geometric information as the baseline is very small compared to the distance to the surface.

In this paper we show that a dense surface geometry estimate can be made by combining the information from a coarse but highly accurate grid of height field points from Laser altimetry measurements and the limited geometrical information from a set of optical images. The resulting surface estimate (both geometry and albedo) has a precision unavailable from other interpolation methods. At the same time far fewer images are needed for a surface reconstruction than without the data

from the laser altimetry measurements.

The calculation is a two step process: Using the images and a spline interpolation of the laser altimetry data, an approximate albedo field of the surface is inferred. This albedo field and the spline interpolated surface are the starting points for the Bayesian surface reconstruction. The varying accuracy of the height field points is taken into account by assigning different uncertainty values to the individual points. The uncertainty of the laser altimetry measured points is very much lower compared with the interpolated values. This approach also offers an easy way to combine measurements with different accuracy. The height field points between the grid points of the laser altimetry measurements are updated by the additional geometrical constraints of the optical images. We present results of the inference of surface models from simulated height field grids and aerial photographs. The influence of different number of images and varying grid resolution is shown.

2 THEORY

The objective here is to infer a *surface model* using the available data, in this case, laser altimeter measurements and optical images. Bayesian inference has, for some time now, been the method of choice for many inference problems, enabling accurate estimation of parameters of interest from noisy and incomplete data (Bernardo 1994). It also provides a consistent framework for the incorporation of multiple, distinct, data sets into the inference process. The general approach is illustrated in figure 1. The figure shows that synthetic observations of the model are made using a computer simulation of the observation process, and that these are compared with the actual observations. The error between the actual and the simulated observations is used to adjust the parameters of the model, to minimize the errors. Bayes theorem tells us directly how much weight to assign to the two sources of errors, those coming from the image measurements and those coming from the laser altimeter measurements.

The surface model we use here is a triangulated mesh. At each vertex of the mesh we store the height and the albedo. As discussed above, to be able to infer the surface heights and albedos, we must first be able to simulate the data that would be recorded from the surface.

Generating images from the surface model is the area of computer graphics known as *rendering* (Foley 1990). It is important to note, however, that much recent work in computer graphics is unsuitable for our purpose, as it works in *image space*, where the fundamental unit is the image pixel, and any given pixel is coloured by light from one and only one surface element. This results in artefacts due to the relative sizes of the projections of the surface elements onto the image plane and their discretization into pixels. These artefacts are particularly noticeable along the edges of the surface elements (aliasing). For this work we require the renderer to operate in *object space*, and below we will briefly describe such a system. We also note that an object-space renderer can also compute the *derivatives* of the pixel values with respect to the surface model parameters. This is crucial in enabling efficient estimation of the surface model parameters, and will be described in more detail below.

We are also required to produce synthetic laser altimeter measurements. We make the approximation that the laser altimeter makes point measurements of the surface, and so producing these synthetic observations is straightforward. We also assume that the error in these measurements are known.

3 A BAYESIAN FRAMEWORK

In this paper the surface geometry is represented by a triangular mesh and the surface reflectance properties (albedos) are associated with the vertices of the triangular mesh. We will consider the case of Lambertian surfaces. We will also assume that the camera parameters (position and orientation, and internal calibration) and the parameters of the lighting are known. It is possible to estimate these parameters in a similar Bayesian framework, but it is beyond the scope of this paper (Morris 2001, Smelyanskiy 2001).

Thus we represent the surface model by the pair of vectors $[\vec{z}\ \vec{\rho}]$. The components of these vectors correspond to the height and albedo values defined on a regular grid of points

$$[\vec{z}\ \vec{\rho}] = \{(z_{\mathbf{i}}, \rho_{\mathbf{i}}), \quad \mathbf{i} = \ell\ (q\,\hat{\mathbf{x}} + p\,\hat{\mathbf{y}})\} \quad q, p = 0, 1, \ldots \quad (1)$$

where ℓ is the elementary grid length, $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ are an orthonormal pair of unit vectors in the (x,y) plane and $\mathbf{i}$ indexes the position in the grid. The pair of vectors of heights and albedos represents a full vector for the surface model

$$\mathbf{u} = [\vec{z}\ \vec{\rho}]. \quad (2)$$

To estimate the values of $\vec{z}, \vec{\rho}$ from the laser altimeter and image data, we apply Bayes theorem which gives

$$p(\vec{z}, \vec{\rho}|L, I_1 \ldots I_F) \propto p(L, I_1 \ldots I_F|\vec{z}, \vec{\rho})\, p(\vec{z}, \vec{\rho}), \quad (3)$$

where L is the laser altimeter data I_f $(f = 1, \ldots, F)$ is the image data. This states that the posterior distribution of the heights and the albedos is proportional to the likelihood – the probability of observing the data given the heights and albedos – multiplied by the prior distribution on the heights and albedos.

Given the surface description, the images and the laser altimeter measurements are *conditionally independent*, and equation 3 can be written as

$$p(\vec{z}, \vec{\rho}|L, I_1 \ldots I_F) \propto p(L|\vec{z}, \vec{\rho})\, p(I_1 \ldots I_F|\vec{z}, \vec{\rho})\, p(\vec{z}, \vec{\rho}),$$

where we now have two independent likelihood terms, one for each data stream.

The prior distribution is assumed to be Gaussian

$$p(\vec{z}, \vec{\rho}) \propto \exp\left(-\frac{1}{2}\mathbf{u}\Sigma^{-1}\,\mathbf{u}^T\right), \quad (4)$$
$$\Sigma^{-1} = \begin{bmatrix} \hat{Q}/\sigma_h^2 & 0 \\ 0 & \hat{Q}/\sigma_\rho^2 \end{bmatrix},$$

where the vector of the surface model parameters $\mathbf{u}$ is defined in (2). The inverse covariance matrix is constructed to enforce a smoothing constraint on local variations of heights and albedos. We penalize the integral over the surface of the curvature factor $c(x, y) = z_{xx}^2 + z_{yy}^2 + 2z_{xy}^2$, and similarly for albedos. The two hyperparameters σ_h and σ_ρ in equation (4) control the expected values of the surface-averaged curvatures for heights and albedos.

This prior is placed directly over the height variables, z, but albedos are only defined over the range $[0-1]$. To avoid this,

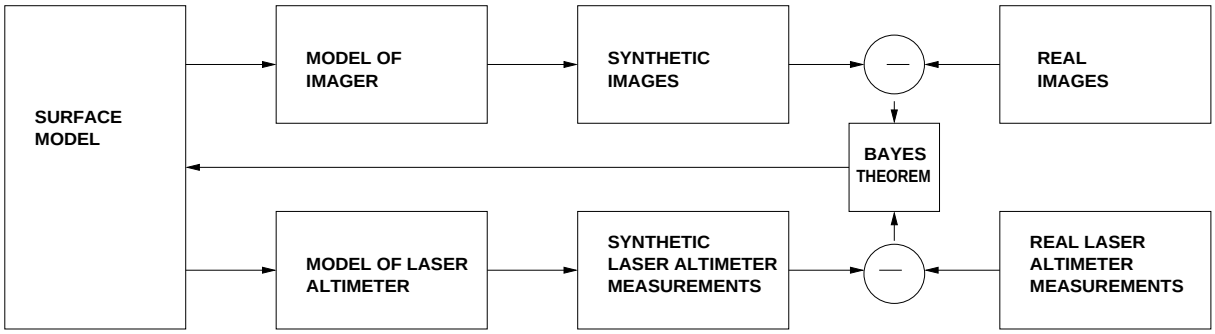

Figure 1: Outline of the Bayesian approach to surface reconstruction from images and laser altimeter measurements

we use transformed albedos $\rho'_{\mathbf{i}}$ in the Gaussian (4), where $\rho'_{\mathbf{i}}$ are defined by:

$$\rho'_{\mathbf{i}} = \log(\rho_{\mathbf{i}}/(1-\rho_{\mathbf{i}})), \quad \mathbf{u} \rightarrow [\vec{z}\ \vec{\rho'}]. \tag{5}$$

In the vector of model parameters $\mathbf{u}$ values of $\vec{\rho}$ are replaced by values of $\vec{\rho'}$.

For both the likelihoods we make the usual assumption that the differences between the observed data and the data synthesized from the model have a zero mean, Gaussian distribution. So for the laser altimeter measurements we have

$$p(L|\vec{z},\vec{\rho}) \propto \exp\left(-\frac{\sum_l (L_l - \hat{L}_l(\vec{z},\vec{\rho}))^2}{2\sigma_l^2}\right) \tag{6}$$

where the summation is over the individual measurement points L_l, and $\hat{L}_l(\vec{z},\vec{\rho})$ denotes the laser altimeter measurements synthesized from the model. The parameter σ_l^2 is the variance of the laser altimeter measurement system.

We also assume that the images I_f comprising the data are conditionally independent, giving

$$p(I_1 \ldots I_F|\vec{z},\vec{\rho}) \propto \exp\left(-\frac{\sum_{f,p}(I_{f\,p} - \hat{I}_{f\,p}(\vec{z},\vec{\rho}))^2}{2\sigma_e^2}\right)$$

where $\hat{I}_{f\,p}(\vec{z},\vec{\rho})$ denotes the pixel intensities in the image f synthesized from the model, σ_e^2 is the noise variance and the summation is over the pixels (p) and over all images (f) used for the inference.

Consider the negative log-posterior.

$$\begin{aligned}\mathcal{L}(\vec{z},\vec{\rho}) \quad \propto \quad & \frac{\sum_{f,p}(I_{f\,p} - \hat{I}_{f\,p}(\vec{z},\vec{\rho}))^2}{\sigma_e^2} \\ & + \frac{\sum_l (L_l - \hat{L}_l(\vec{z},\vec{\rho}))^2}{\sigma_l^2} \\ & + \mathbf{x}\,\Sigma^{-1}\,\mathbf{x}^T,\end{aligned} \tag{7}$$

where $\mathbf{x} = \mathbf{u} - \mathbf{u}_0$ is a deviation from a current estimate $\mathbf{u}_0$. $\mathcal{L}$ is a nonlinear function of $\vec{z},\vec{\rho}$ and the MAP estimate is that value of $\vec{z},\vec{\rho}$ which minimizes $\mathcal{L}(\vec{z},\vec{\rho})$.

The crux of the problem is thus how to minimize $\mathcal{L}$. We apply a gradient method, using an initialization based on a spline interpolation of the laser altimeter measurements.

Making the assumption that the laser altimeter makes point measurements of one of the vertices of the mesh, equation 6 can be written as

$$p(L|\vec{z},\vec{\rho}) \propto \exp\left(-1/2(\mathbf{l}-\mathbf{l}_0)\Sigma_l^{-1}(\mathbf{l}-\mathbf{l}_0)^T\right)$$

where $\mathbf{l}_0$ is the vector of actual laser altimeter observations, and $\mathbf{l}$ are the corresponding entries taken from the $\vec{z}$ vector. The inverse covariance matrix σ_l^{-1} is a diagonal matrix with $1/\sigma_l^2$ on the leading diagonal.

The term for the image measurements is more complex, as $\hat{I}(\vec{z},\vec{\rho})$ is the rendering process. To make progress with minimizing $\mathcal{L}(\vec{z},\vec{\rho})$ we linearize $\hat{I}(\vec{z},\vec{\rho})$ about an initial estimate, $\vec{z}_0,\vec{\rho}_0$

$$\hat{I}(\vec{z},\vec{\rho}) = \hat{I}(\vec{z}_0,\vec{\rho}_0) + \mathbf{D}\,\mathbf{x}, \quad \mathbf{D} \equiv \left\{\frac{\partial \hat{I}_{f\,p}}{\partial z_{\mathbf{i}}}, \frac{\partial \hat{I}_{f\,p}}{\partial \rho'_{\mathbf{i}}}\right\} \tag{8}$$

where $\mathbf{D}$ is the matrix of derivatives evaluated at z_0,ρ_0. Then the minimization of $L(\vec{z},\vec{\rho})$ is replaced by minimization of the quadratic form:

$$\begin{aligned}\mathcal{L}' &= \frac{1}{2}\mathbf{x}\,\hat{A}\,\mathbf{x} - \mathbf{b}\,\mathbf{x}, \quad \mathbf{x} \equiv \mathbf{u} - \mathbf{u}_0, \\ \hat{A} &= \Sigma^{-1} + \frac{\mathbf{D}\mathbf{D}^T}{\sigma_e^2} + \hat{\Sigma}_l^{-1}, \\ \mathbf{b} &= \frac{(I - \hat{I}(\vec{z}_0,\vec{\rho}_0))}{\sigma_e^2}\mathbf{D}\end{aligned} \tag{9}$$

where $\hat{\Sigma}_l^{-1}$ is now a large square matrix (of dimension length($\vec{z}$) + length($\vec{\rho}$)), where the diagonal elements corresponding to the vertices for which there are laser altimeter measurements take values $1/\sigma_l^2$ and all other entries are zero. The entries of $\mathbf{u_0}$ corresponding to the laser altimeter measurements are set to the observed values, and the remaining height values are initialized using a spline interpolation. This interpolation and the albedo initialization will be described below.

In equation 9 $\hat{A}$ is the Hessian matrix of the quadratic form and vector $\mathbf{b}$ is the gradient of the likelihood $\mathcal{L}$ computed at the current estimate. We search for the minimum in $\mathbf{x}$ using a conjugate-gradient method (Press 1992).

Thus the most difficult part of finding the MAP is the requirement to render the image and compute the derivatives for any values of the surface model parameters. We discuss this computation in some detail in the next sections. Here it is sufficient to note that while forming $\hat{I}$ using only object space computation (see section 4) is computationally expensive, we can compute $\mathbf{D}$ at the same time for little additional computation. Also the derivative matrix is sparse with the number of nonzero entries a few times the number of model parameters. This makes the process described above a practical one.

The log-posterior is potentially multi-modal, and so it is important to begin the optimization from a good initialization.

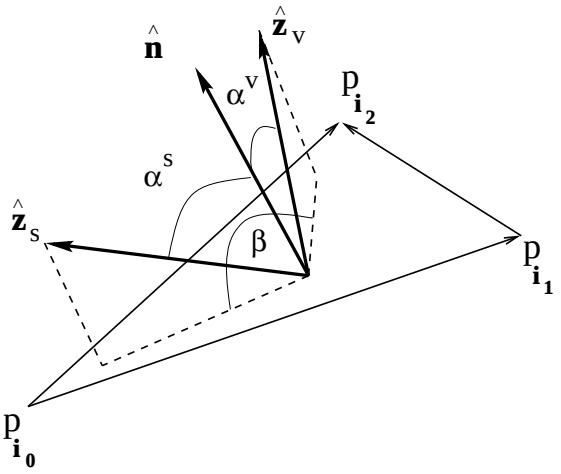

Figure 2: Geometry of the triangular facet, illumination direction and viewing direction. $\hat{\mathbf{z}}_s$ is the vector to the illumination source; $\hat{\mathbf{z}}_v$ is the viewing direction.

In order to do this, the high resolution surface estimation proceeds as follows.

1. Use a spline interpolation of the laser altimeter measurements to produce an initial height field estimate at the desired high resolution.
   Also produce the matrix $\hat{\Sigma}_l^{-1}$, with zeros everywhere except those diagonal entries corresponding to the points of the high resolution surface grid that are measured by the laser altimeter. These values are $1/\sigma_e^2$.
   Initialise all the albedos to 0.5.
2. Render the surface generated in step 1 and compute the derivative matrices **D** (one for each image). Set to zero all the derivatives with respect to the surface heights. This fixes the heights at their current values in the optimization step (below), so that only the surface albedo values are inferred.
   Starting from the surface from step 1, and using the derivative matrices calculated above, use the conjugate gradient algorithm to minimize the linearization of the log-posterior in equation 9. This produces a good initialization for the final optimization.
3. Render the surface generated in step 2 and compute the derivative matrices **D**.
   Starting from the surface in step 2, use the conjugate gradient algorithm to minimize the linearization of the log-posterior.
4. Repeat step 3 until convergence.
   The minimum found is the final surface estimate, which combines the information from the laser altimeter measurements and the visible images.

4 FORMATION OF THE IMAGE AND THE DERIVATIVE MATRIX.

The task of forming an image, $\hat{I}$, given a surface description, $\vec{z}, \vec{\rho}$, and camera and illumination parameters is the area of computer graphics known as rendering (Foley 1990). Most current rendering technology is focused on producing images which are *visually* appealing, and producing them very quickly. As discussed in the introduction, this results in the use of image-space algorithms, with the fundamental assumption that each triangle making up the surface, when projected onto the image plane, is much larger than a pixel. This makes reasonable the assumption that any given pixel receives light from only one triangle, but does produce images with artifacts at the triangle edges. Standard rendering also produces inaccurate images if the triangles project into areas much smaller than a pixel on the image plane, as the pixel will then be colored with a value coming from just *one* of the triangles.

Clearly this approach is not suitable for high-resolution 3D surface reconstruction from multiple images. The triangles in a high-resolution surface may project onto an area much smaller than a single pixel in the image plane (sub-pixel resolution). Therefore, as discussed in the introduction, for our system we implemented a renderer for triangular meshes which performs all computation in *object space*. At present we neglect the blurring effect due to diffraction and due to the role of pixel boundaries in the CCD array. Then the light from a triangle as it is projected into a pixel contributes to the brightness of the pixel with a weight factor proportional to the fraction of the area of the triangle which projects into that pixel. This produces anti-aliased images and allows an image of any resolution to be produced from a mesh of arbitrary density, as required when the system performing the surface inference may have no control over the image data gathering.

Our renderer computes brightness $\hat{I}_p$ of a pixel p in the image as a sum of contributions from individual surface triangles $\triangle$ whose projections into the image plane overlap, at least partially, with the pixel p.

$$\hat{I}_p = \sum_{\triangle} f_{\triangle}^p \, \Phi_{\triangle}. \tag{10}$$

Here $\Phi_{\triangle}$ is a radiation flux reflected from the triangular facet $\triangle$ and received by the camera, and $f_{\triangle}^p$ is the fraction of the flux that falls onto a given pixel p in the image plane. In the case of Lambertian surfaces and a single spectral band $\Phi_{\triangle}$ is given by the expression

$$\begin{aligned} &\Phi_{\triangle} = \rho \, E(\alpha^s) \, \cos\alpha^v \, \cos^{\kappa}\theta \, \Delta\Omega, \\ &E(\alpha^s) = \mathcal{A}\,(\mathcal{I}^s \, \cos\alpha^s + \mathcal{I}^a)\,. \\ &\Delta\Omega = S/d^2. \end{aligned} \tag{11}$$

Here ρ is an average albedo of the triangular facet. Orientation angles α^s and α^v are defined in figure 2. $E(\alpha^s)$ is the total radiation flux incident on the triangular facet with area $\mathcal{A}$. This flux is modeled as a sum of two terms. The first term corresponds to direct radiation with intensity $\mathcal{I}^s$ from the light source at infinity (commonly the sun). The second term corresponds to ambient light with intensity $\mathcal{I}^a$. The parameter θ in equation. (11) is the angle between the camera axis and the viewing direction (the vector from the surface to the camera); The exponent κ is the lens falloff factor. $\Delta\Omega$ in (11) is the spatial angle subtended by the camera which is determined by the area of the lens S and the distance d from the centroid of the triangular facet to the camera.

We identify the triangular facet $\triangle$ by the set of 3 indices $(\mathbf{i_0}, \mathbf{i_1}, \mathbf{i_2})$ from the vector of heights (1) that determines the vertices of the triangle in a counterclockwise direction (see figure 2). In the r.h.s of equation (11) we have omitted for brevity those indices from all the quantities associated with individual triangles. The average value of albedo for the

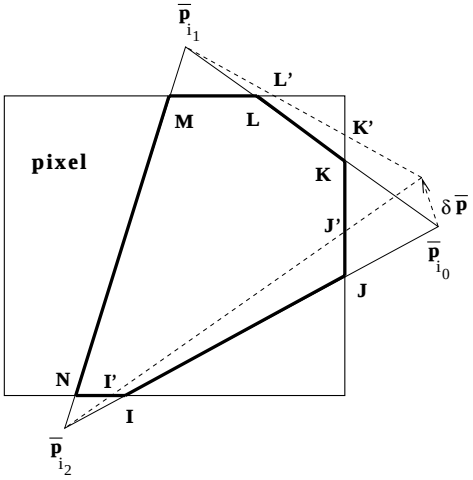

Figure 3: The intersection of the projection of a triangular surface element $(\mathbf{i_0}, \mathbf{i_1}, \mathbf{i_2})$ onto the pixel plane with the pixel boundaries. Bold lines corresponds to the edges of the polygon resulting from the intersection. Dashed lines correspond to the new positions of the triangle edges when point $\mathbf{P_{i_0}}$ is displaced by $\delta\mathbf{P}$

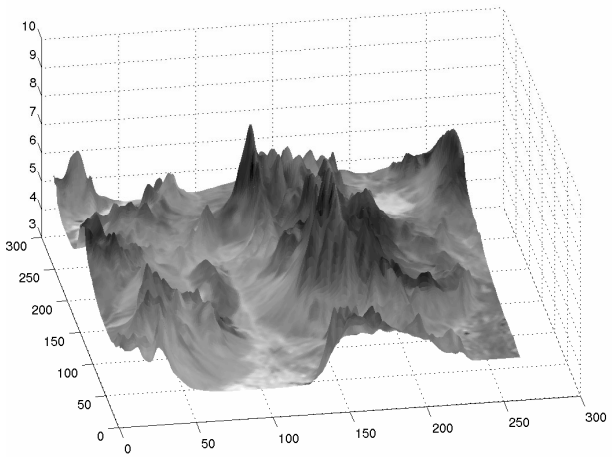

Figure 4: The initial synthetic surface model (Duckwater, NV). (Note the exaggerated vertical scale.)

triangle in (11) is computed based on the components of the albedo vector ρ corresponding to the triangle indices

$$\rho_{\triangle} \equiv \rho_{\mathbf{i_0},\mathbf{i_1},\mathbf{i_2}} = \frac{1}{3}(\rho_{\mathbf{i_0}} + \rho_{\mathbf{i_1}} + \rho_{\mathbf{i_2}}). \tag{12}$$

We note that using average albedo (12) in the expression for $\Phi_{\triangle}$ is an approximation which is justified when the albedo values vary smoothly between the neighboring vertices of a grid.

The area $\mathcal{A}$ of the triangle and the orientation angles in (11) can be calculated in terms of the vertices of the triangle $\mathbf{P_i}$ (see figure 2) as follows:

$$\hat{\mathbf{n}} \cdot \hat{\mathbf{z}}^s = \cos\alpha^s, \quad \hat{\mathbf{n}} \cdot \hat{\mathbf{z}}^v = \cos\alpha^v, \tag{13}$$
$$\hat{\mathbf{n}} = \frac{\mathbf{v_{i_0,i_1}} \times \mathbf{v_{i_1,i_2}}}{2\mathcal{A}}, \quad \mathbf{v_{i,j}} = \mathbf{P_j} - \mathbf{P_i}$$

Here $\hat{\mathbf{n}}$ is a unit normal to the triangular facet and vectors of the edges of the triangle $\mathbf{v_{i,j}}$ are shown in figure 2.

We use a standard pinhole camera model with no distortion in which coordinates of a 3D world point $\mathbf{P} = (x, y, z)$ are first rotated with the rotation matrix $\hat{\mathbf{R}}$ and then translated by the vector $\mathbf{T}$ into camera coordinates, yielding $\mathbf{P_c} = (x_c, y_c, z_c)$

$$\mathbf{P_c} = \hat{\mathbf{R}}\,\mathbf{P} + \mathbf{T} \tag{14}$$

($\hat{\mathbf{R}}$ and $\mathbf{T}$ are expressed in terms of the camera registration parameters (Hartley 2000). We do not give them explicitly here). After the 3D transformation given in (14), point $\mathbf{P_c}$ in the camera coordinate system is transformed using a perspective projection into the 2D image point $\bar{\mathbf{P}} = (\bar{x}, \bar{y})$ using a focal length f and aspect ratio a.

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = -\frac{f}{z_c} \begin{bmatrix} a\,x_c \\ y_c \end{bmatrix}. \tag{15}$$

We use 2D image projections of the triangular vertices $\mathbf{P_i}$ to compute the area fraction factors $f^p_{\triangle}$ for the surface triangles (cf. Eq. (10))

$$f^p_{\triangle} = \frac{\bar{A}_{\text{polygon}}}{\bar{A}_{\triangle}}. \tag{16}$$

Here $\bar{A}_{\triangle}$ is the area of the projected triangle on the image plane and $\bar{A}_{\text{polygon}}$ is the area of the polygon resulting from the intersection of the projected triangle and the pixel p (see figure 3).

4.1 Computation of the derivative matrix.

The inference of the surface model parameters depends on the ability to compute the derivatives of the modeled observations $\hat{I}$ with respect to the model parameters. According to equation (10), the intensity $\hat{I}_p$ of a pixel p depends on the subset of the surface parameters, (heights and albedos), that are associated with the triangles whose projections overlap the pixel area.

The derivatives $\hat{I}_p$ with respect to logarithmically transformed albedo values are easily derived from equations (5), (10) and (11).

In our object-space renderer, which is based on pixel-triangle geometrical intersection in the image plane, the pixel intensity derivatives with respect to the surface heights have two distinct contributions

$$\frac{\partial \hat{I}_p}{\partial z_{\mathbf{i}}} = \sum_{\triangle} \left(f^p_{\triangle}\ \frac{\partial \Phi_{\triangle}}{\partial z_{\mathbf{i}}} + \Phi_{\triangle} \frac{\partial f^p_{\triangle}}{\partial z_{\mathbf{i}}} \right) \tag{17}$$

Variation of the surface height $z_{\mathbf{i}}$ gives rise to variations in the normals of the triangles associated with this height (in a general triangular mesh, on average 6 triangles are associated with each height) and this produces the derivatives of the total radiation flux $\Phi_{\triangle}$ to the camera from those triangles. This is the first term in equation (17). Also, height variation gives rise to the displacement of the corresponding point which is the projection of this vertex on the image plane. This results in changes to the areas of the triangles and polygons with edges containing this point (see figure 3). This produces the derivatives of the fractions $f^p_{\triangle}$, the second term in equation 17. Details of these derivatives can be found in (Smelyanskiy 2000, Morris 2001, Smelyanskiy 2001).

5 RESULTS

Figure 4 shows the synthetic surface that we will use to demonstrate our methodology. The topography is taken from the USGS DEM of Duckwater, Nevada. A LANDSAT-TM

image 1	camera	$(75, 150, 2000)$
	look at	$(150, 150, 0)$
	view up	$(0, 1, 0)$
image 2	camera	$(225, 150, 2000)$
	look at	$(150, 150, 0)$
	view up	$(0, 1, 0)$

Table 1: Camera parameters used to generate the images in figure 5

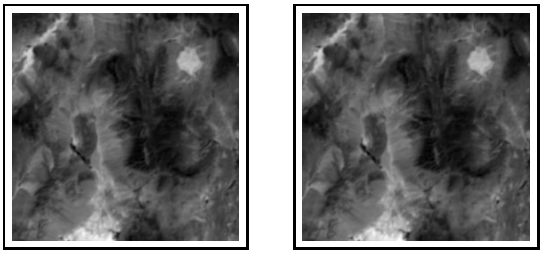

Figure 5: Images of the synthetic surface

image was co-registered with the DEM, and the values of one band were used in place of the true albedos. This results in the surface shown. One unit is approximately 180 meters.

Figure 5 shows two images rendered from the surface, and table 1 gives the positions and orientations of the (synthetic) cameras. The cameras were positioned to approximate satellite observations. The two images look very realistic. Note that the images appear very similar due to the proximity of the two camera positions. There is limited geometric information available from the images alone.

Figure 6 shows a surface from a grid of 9×9 points extracted from the surface in figure 4. The major terrain features have all been sampled, but clearly it is a very poor representation of the surface. This is taken as the laser altimeter observations of the surface.

Using the images and the 9×9 grid, we will now go through the surface estimation procedure that was detailed above.

Figure 7 shows the result of using the standard spline interpolation to expand the 9×9 grid to the full resolution of the surface. The result is a smooth surface showing the major features, but note that it contains no more information than the coarse surface.

Keeping the heights fixed at this surface, we then use the images to infer initial values for the albedos. The result is shown in figure 8. Note that this is *not* simply the back-projection of the of the images onto the surface, The information in both images has been optimally combined to give the albedo estimates. This surface is now a passable approximation to the original surface, as it has high resolution albedo information providing rich visual detail, but clearly it contains no topographic detail.

Figure 9 shows the final inferred surface. This is much improved over the surface in figure 8. It shows that much of the detail of the topography has been extracted from the data and incorporated into the model. The error surfaces show in figures 10 and 11 show clearly the improvement in the surface estimate. These error surfaces show both the height error and the albedo error as a shaded surface – the topography shows the height error, and the colour of the surface

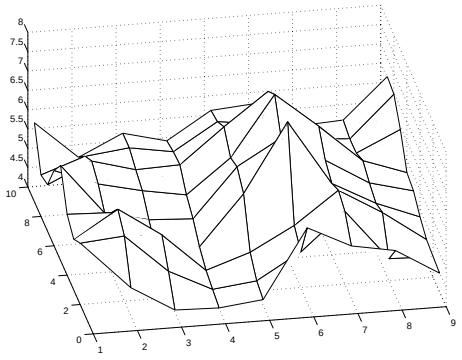

Figure 6: Grid from the 9×9 simulated laser altimeter measurements.

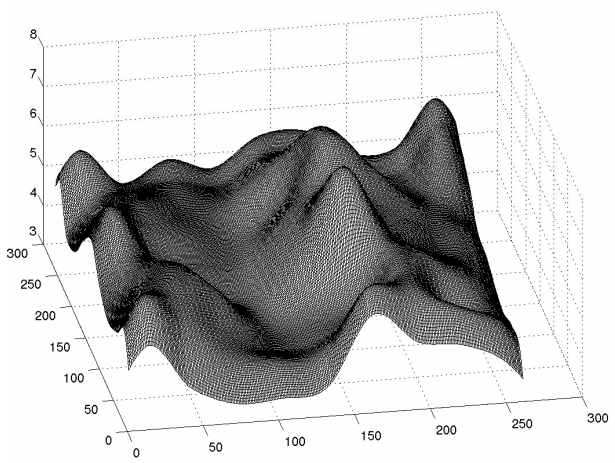

Figure 7: Spline interpolation of the synthetic laser altimeter measurements.

shows the albedo error. The rms errors for the interpolated surface are 0.64 (per vertex) for the heights and 2.6×10^{-5} for the albedos. For the final infered surface the errors are 0.01 for the heights and 1.3×10^{-5} for the albedos. Note that the albedo values from the initialization are already quite good (as can be seen on figure 8), however the inference process produces a topography which is very significantly more accurate.

6 CONCLUSIONS AND FUTURE EXTENSIONS

We have presented the theory and practice of using Bayesian methodology to combine the information in laser altimeter measurements and visible images into a single, high resolution surface model. We have shown on synthetic data that the two data sets can be combined into a single high resolution model that is more detailed than could be provided by either data stream alone.

Current work is proceeding towards applying the demonstration system to real data, including NASA mission data. Work in this area is devoted to sensor modeling (producing the synthetic images and derivative matrices for the actual imaging sensor, rather than an idealization of it), estimation of the camera positions to sub-pixel accuracy, better control of the smoothness prior on the surface, and better initialization of the optimization procedure.

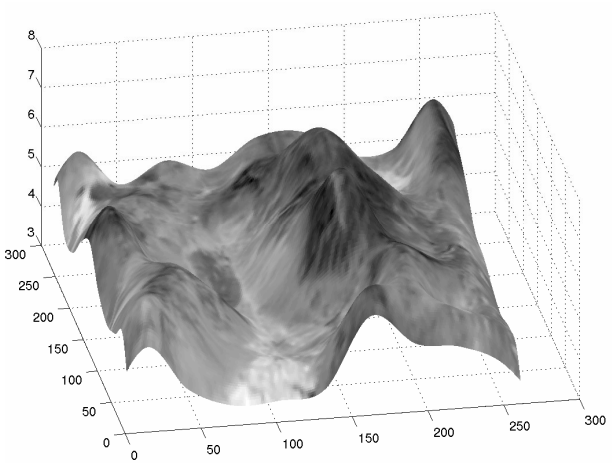

Figure 8: Interpolated surface with inferred albedos.

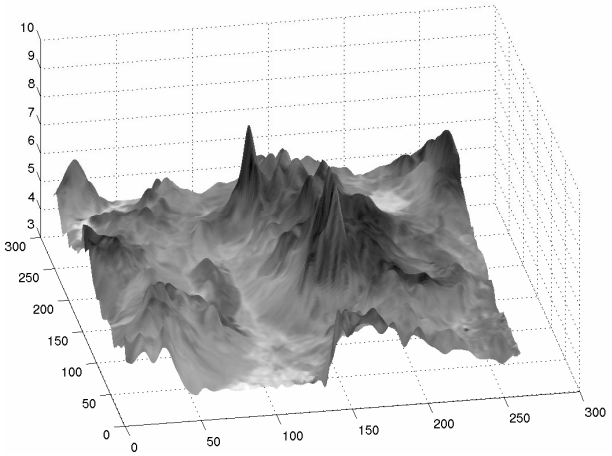

Figure 9: Final inferred surface using the Bayesian approach to combining the laser altimeter and visible images

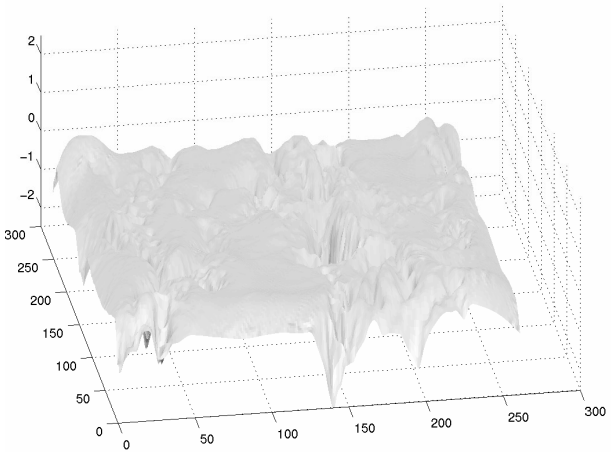

Figure 10: Error surface for the interpolated surface

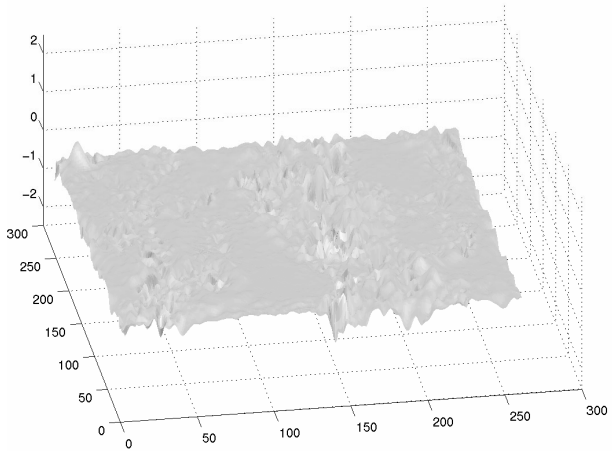

Figure 11: Error surface for the inferred surface

REFERENCES

Bernardo, J. and Smith, A., 1994. Bayesian Theory. Wiley, Chichester, New York.

Foley, J., van Dam, A., Finer, S. and Hughes, J., 1990. Computer Graphics, Principles and Practice. Addison-Wesley, 2nd edition.

Hartley, R. and Zisserman, A., 2000. Multiple View Geometry in Computer Vision. Cambridge University Press.

Morris, R.D., Smelyanskiy, V.N., Cheeseman, P., 2001. Matching Images to Models – Camera Calibration for 3-D Surface Reconstruction. Proceedings of EMMCVPR 2001, Springer LNCS volume 2134.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical Recipes in C. Cambridge University Press, 2nd edition.

Rees, W., 1990. Physical Principles of Remote Sensing. Cambridge University Press.

Smelyanskiy, V.N., Cheeseman, P., Maluf, D.A., Morris, R.D., Bayesian Super-Resolved Surface Reconstruction from Images. Proceedings of Computer Vision and Pattern Recognition 2000.

Smelyanskiy, V.N., Morris, R.D., Maluf, D.A., Cheeseman, P., 2001. (Almost) Featureless Stereo – Calibration and Dense 3D Reconstruction Using Whole Image Operations. Technical Report, RIACS, NASA Ames Research Center.